Classifying Skin Lesion Images Using Adaptive Boosting

Miriam Hu[1], Catherine Myong[2], and Ben Sadis[3]
[1]Indiana University, [2]Harvard University, [3]University of Michigan

Background

Melanoma:
- The most serious form of skin cancer
- Responsible for more than 125,000 deaths in the United States from 1999 to 2013[1]
- Early detection can dramatically increase the survival rate of melanoma:
  - » 50 percent in late stages when lesion is ≥ 4 mm in depth
  - » 95 percent if caught when the lesion is < 1 mm in depth[2]

Diagnosis:
- Physicians use the ABCDEs (asymmetry, border irregularity, color variegation, diameter ≥ 6 mm, and evolution of size/shape/color) to classify a skin lesion as either benign or malignant.
- If the lesion is considered suspicious, the physician will perform a biopsy for a pathologist to analyze it.

Current Limitations:
- Biopsy is costly in terms of time and money.
- Dermatologists do not always classify skin lesions accurately.
- Convolutional neural networks (CNNs) have good accuracy but are computationally intensive[3].

Feature Extraction

Data Set:
- Classification models were trained on a set of 700 photographs of skin lesions (535 benign and 135 malignant) from the International Skin Imaging Collaboration (ISIC).
- The testing set has 200 images from the same source.

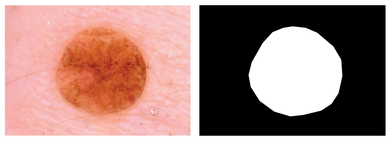


Figure 1: A benign lesion and its binary segmentation mask, where the lesion area is white and skin/background is black.

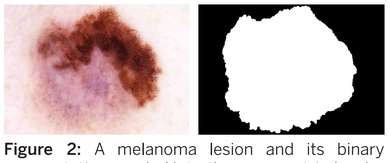


Figure 2: A melanoma lesion and its binary segmentation mask. Note the asymmetric border and varied colors within the lesion as compared to the benign lesion.

Texture Features:
- Extracted from the grayscale images using R's `radiomics` package
- First-order features: Features only depending on the pixels themselves
- Second-order features: features depending on the relationship between two or more pixels
  Ex: The gray level co-occurrence matrix shows how many neighboring pixels have the same gray level and in general measures the smoothness of the image.

Color Features[4]:
- Extracted from the RGB (red, green, blue) images
- Mean and standard deviation of each color channel
- Mean and standard deviation of the intensity channel (**Equation 2**)
- Color variegation (**Equation 3**) for each color channel and for the intensity

$$I(x,y) = \sqrt{R^2(x,y) + G^2(x,y) + B^2(x,y)}$$

Equation 1: Definition of the intensity channel for each pixel I(x, y) in the image.

$$cv_R = log\frac{\sigma(R)}{\mu(R)}$$

Equation 2(a): Parameter for the color variegation of the red channel. (Calculation for green and blue channels is analogous.)

$$cv_I = log\frac{\sigma(I)}{\mu(I)}$$

Equation 2(b): Parameter for the color variegation of the intensity channel.

Asymmetry and Circularity[5]:
- According to the ABCDE principles, lesions that are more asymmetric (**Equation 3**) should be more likely to be malignant.
- Lesions that are more circular (**Equation 4**) should be less likely to be malignant.

$$SD = \frac{1}{n}\sum_{i=0}^{n-1}\|Pi - \widehat{Pi}\|^2$$

Equation 3: Measure of asymmetry. The symmetry distance (SD) is the displacement of each vertex when the skin lesion is transformed into a symmetric shape.

$$circ = \frac{4\pi A}{P^2}$$

Equation 4: Measure of circularity as a function of area *A* and perimeter *P*. Objects with ratio closer to 1 are more circular, while objects with ratio closer to 0 are less circular.

Model Creation and Selection

Diagnostic Measures:

$$TPR = \frac{TP}{P} \; or \; \frac{TP}{TP+FN}$$

Equation 5(a): Sensitivity: The true positive rate (TPR), or percentage of malignant lesions that are correctly identified as malignant.

$$TNR = \frac{TN}{N} \; or \; \frac{TN}{TN+FP}$$

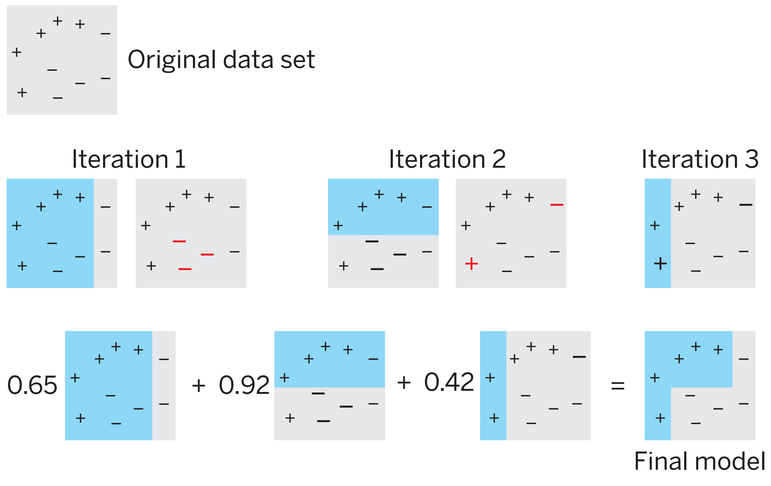
Equation 5(b): Specificity: The true negative rate (TNR), or percentage of benign lesions that are correctly identified as *not* being malignant.

$$avg\ acc = \frac{TPR+TNR}{2}$$

Equation 5(c): Balanced (average) accuracy: Measures how well the classifier correctly identifies benign and malignant cases.

AdaBoost: "Adaptive Boosting"
- Fit using `ada` package in R
- Uses many weak learners (bad at classification) to create a strong learner (good at classification)
- Iterative process
- When an image is misclassified, it is given more weight in the next iteration[6].

Tuning:
- Choose optimal parameters:
  - » Exponential loss function
  - » Number of iterations = 50
- Select best feature set:
  - » First-order features
  - » Color features
  - » Circularity
  - » Including more features makes model performance worse.



Predictions \ Truth	*Benign*	*Malignant*
Benign	112	10
Malignant	50	28

Figure 3: Confusion matrix for testing set using AdaBoost model. Correct predictions are along the diagonal.

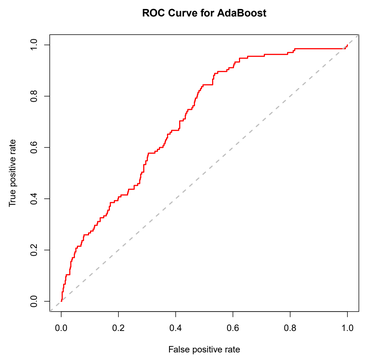


Figure 2: Receiver operating characteristic (ROC) curve for AdaBoost model based on 10-fold cross validation repeated 10 times. Ideally, the curve should appear hyperbolic (indicates high sensitivity and specificity).

Conclusion

We fit several models to classify skin lesions as benign or malignant:
- Support vector machine (SVM)
- Random forest
- Neural network
- AdaBoost

The AdaBoost model performed best.
Sensitivity: 0.736
Specificity: 0.691
Average accuracy: 0.714

Dermatologists achieved about 65.8 percent accuracy[3].

References

[1]U.S. Cancer Statistics Working Group. United States Cancer Statistics: 1999–2014. Incidence and Mortality Web-based Report. Atlanta (GA): Department of Health and Human Services, Centers for Disease Control and Prevention, and National Cancer Institute.
[2]Diagnosis and Treatment of Early Melanoma. NIH Consens. Statement 1992;10(1):1–26.
[3]Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115–118.
[4]Jaworek-Korjakowska J. Computer-Aided Diagnosis of Micro-Malignant Melanoma Lesions Applying Support Vector Machines. BioMed Research International. 2016;2016.
[5]Ng V, Cheung D. Measuring asymmetries of skin lesions. Proceedings of the IEEE International Conference on Systems, Man and Cybernetics. 1997;5:4211–4216.
[6]Schapire RE. Explaining AdaBoost. Empirical Inference. Springer, Berlin, Heidelberg. 2013.

Acknowledgments

We would like to thank Dr. Tim Johnson, Dr. Jian Kang, Dr. Eunjee Lee, and Cui Guo for mentoring the Imaging group and answering all our questions. We would also like to thank Dr. Bhramar Mukherjee, Mitch Sevigny, Davina Barron, Sabrina Clayton, and everyone else who helped make the Big Data Summer Institute a wonderful experience this summer.